AUT Journal of Mechanical Engineering

AUT J. Mech. Eng., 1(1) (2017) 109-116
DOI: 10.22060/mej.2016.757

Modeling Behavior of a Chain of Protein as a Bionanorobot by Changing Environmental Condition

A. Gaffari[1], A. Khodayari[2*], Sh. Movasagh[1]

[1] Faculty of Mechanical Engineering, K. N. Toosi University of Technology, Tehran, Iran
[2] Faculty of Mechanical Engineering, Pardis Branch Islamic Azad University, Tehran, Iran

ABSTRACT: In this paper, the behavior of one chain of small heat shock protein based on molecular dynamic simulation is modeled by the nonlinear identifier. Molecular dynamic simulation is time-consuming due to the high load of calculations, and providing a model of system behavior facilitates utilizing protein as a bio nano robot in a shorter time. The ARC1 is a molecular Chaperone with swarm structure containing 12 chains. Each chain of small heat shock protein contains two sections; arm and the central cavity which is introduced as a bionanorobot for their special biological structure and their reaction to external forces. The conformational changes of protein with one chain due to external excitation have been analyzed. In addition to system identification of one chain of small heat shock protein, the effect of temperature, pH, and content of solvent are examined on the behavior of bionanorobot arm and a central cavity in a wide range of variation. The results show that minimum number of error is relevant to the adaptive neuro fuzzy system identifier. Modeling the behavior of one chain provides a suitable condition to control the central cavity and bionanorobot arm in a shorter period of time compared to the molecular dynamic simulation.

Review History:
Received: 19 December 2015
Revised: 8 March 2016
Accepted: 9 May 2016
Available Online: 8 November 2016

Keywords:
Chaperone protein
Bionanorobot
GROMACS
Molecular dynamic
Nonlinear identifier

1- Introduction

Bionanorobotics is an interdisciplinary and emerging field regarding assembly, manufacturing and using biomolecular devices in nanoscale. The most important function of bionano robots is in medical tasks aiming and diagnostics, curing and using in surgery [1-3]. Biological tests of bacterial functions [4] and magnetic systems show that they are nanoscale machines without a help of external hardware. Also, proteins and DNAs are able to act as a motor, mechanical attachments or sensors [5]. The nature of protein is to achieve a certain cellular task, by replacing a bulk to react against the catalyst, but DNA is introduced as a data carrier. Therefore, most molecular machines are comprised of proteins. Advances made in this field is to the extent that such components are used for drug delivery [1, 3, 6] or are used in techniques of engineering agitation [7-9]. The best choices in this regard are nanotubes due to their mechanical and electrical properties, DNA and proteins.

Among proteins, Chaperone has gained interests due to the proper biological structure. In molecular biology, Chaperone is a protein that helps folding, assembly and detachment of macromolecular structures [10]. Many Chaperone proteins are heat shock molecules produced in stressing conditions of the cells. The cells play their roles in hindering changes in the combination and the structure of protein under stressors such as high temperature, pressure or cells tension [11]. This protein exits in all live cells in an attached or detached form to other proteins. These proteins are located in cytoplasm. Perfoldine is one of Chaperone type proteins, which was introduced first by Hassanzadeh et al. [12-14] as a nano-acuator, and by mutating and applying control inputs of temperature and pH showed the ability of nanoactuator for taking and releasing a Nano cargo.

In this article, the behavior of sHSP is considered based on molecular dynamic simulating. Each chain of this protein is introduced as a protein bionano robot. By choosing proper inputs and considering the behavior of the system in single chain mode in a wide range of inputs, the dynamic model of each chain is obtained based on the non-linear identifier. Finally, various criteria of errors are used for validation.

2- Model and Simulation Details

sHSPs are present everywhere and play important role in cell protection, agglomeration of denatured proteins and facilitating folding by other Chaperones [15, 16]. ARC1 Chaperone protein is composed of 12 chains (PDB file code: 2BYU) and has a bulk structure (Fig. 1) [17].

In order to simulate molecular dynamic, we used GROMACS software. The input of this software is PDB file of intended protein. PDB file provides an archive of atomic coordinates and other data describing proteins and important biological macromolecules. According to the similarity of chains for analysis of the behavior of this protein, only one of them is placed in molecular dynamic simulation box. Fig. 2 shows the structure of a chain of sHSP chaperone protein. Each chain is composed of the arm and mass center. During dynamic molecular simulation, by applying variable inputs, displacement of the arm and mass center for a chain of this protein is examined.

Molecular dynamic simulation is a kind of computer simulation method. Molecular systems include numerous atoms and it is

Corresponding author, E-mail: Khodayari@pardisiau.ac.ir

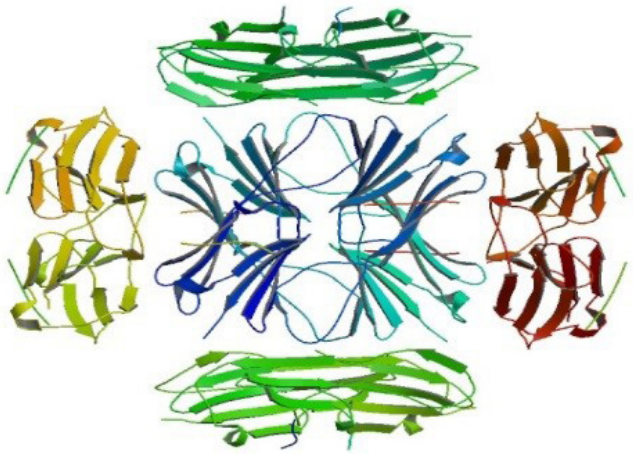

Fig. 1. Schematic structure of sHSP

impossible to obtain some characteristics analytically [18]. GROMACS software is for molecular dynamic simulation. GROMACS is a calculation motor for molecular dynamic simulation and minimizing energy [19]. Choosing the type of force field in molecular dynamic simulation determines potential energy function. For molecular dynamic simulation in GROMACS software, some prerequisites are necessary. In this article, Amber99 force field, a cubic box and its size are determined so that the mass center of protein is 1nm away from the walls of the box. SPC-216 solvent is used and gradient descent method is used to minimize potential function.

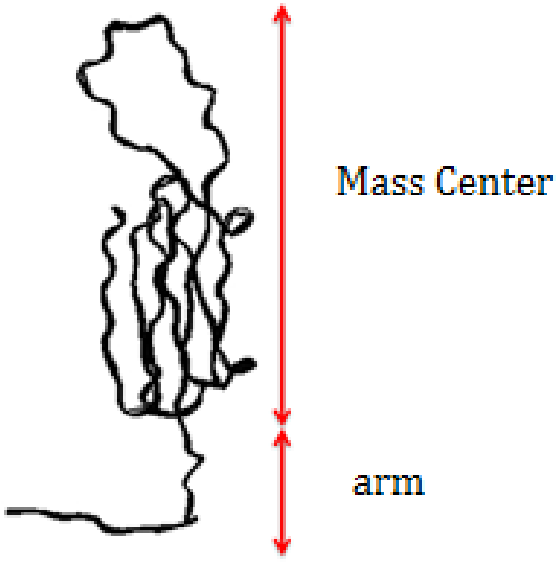

Fig. 2. The structure of a chain from sHSP, obtained by omitting other chains from PDB file

3- 3. Molecular Dynamic Simulation

By molecular dynamic simulation, a chain of sHSP is observed. By changing temperature, pH and change in the size of the box which leads to a change in concentration of solvent, the behavior of mass center and its arm changes significantly. Therefore, change in temperature, pH, and concentration of solvent are considered as three main input variables for Chaperone protein system. The details of the procedure are as follows:

3- 1- Temperature change

First, a chain of sHPS undergoes molecular dynamic simulation in normal condition, i.e., 310 K and null pH. In order to analyze the behavior of the system in the wider range of temperature, by applying descending slope temperature from 310 to 270 K, the behavior of arm and replacing mass center in a single chain mode are examined. Fig. 3 shows descending temperature plot during molecular dynamic simulation.

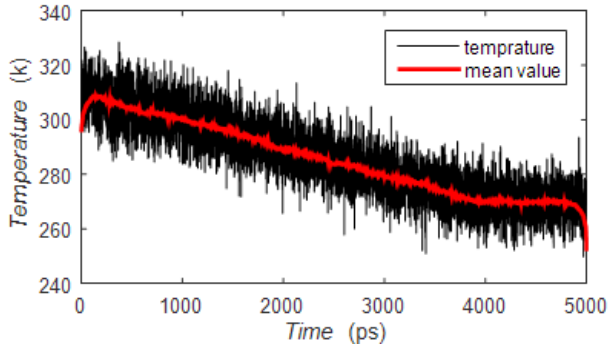

Fig. 3. Decreasing temperature from 310 to 270 K during molecular dynamic simulation

3- 2- pH change

pH changes lead to the change in electrostatic electrical load around protein. In normal mode, pH of solvent around a chain of protein in the molecular dynamic simulation is equal to 7.4. In order to change pH of solvent around protein, a series of amino acids must protonate and others have to deprotonate. Among 20 types of existing amino acids, histidine, glutamic acid, and aspartic acid must protonate in acidic pH, also lysine and Arginine must deprotonate in an alkali environment. In order to provide an acidic environment, pH around protein has been considered equal to 3, and for alkali surrounding, pH has to be set to 11.

To examine the behavior of a chain of sHSP against pH change of solvent, first, pKa of amino acids of proteins have to be calculated. For this purpose, it is appropriate to use Propka server. By calculating pKa and drawing Fig. 4, we can figure it out in molecular dynamic simulation, which residues of each amino acids must gain proton to lose it.

In the acidic mode, the residues with pKa higher than 3 protonate, and in alkali mode, residues with pKa less than 11 deprotonate. Finally, to provide a biological, electrically null surrounding, positive and negative ions of halide are added to the solvent. During pH change, the temperature is constant and set to 310 K.

3- 3- Solvent concentration change

When the molecular dynamic simulation of macromolecules such as protein is done in a certain solvent, usually intermittent boundary conditions must be used to minimize the effect of box size in the simulation. In such simulations, the cost of calculations is related to solvent's degree of freedom. Therefore, size and type of the box will be significantly effective in optimizing the level of solvent [20].

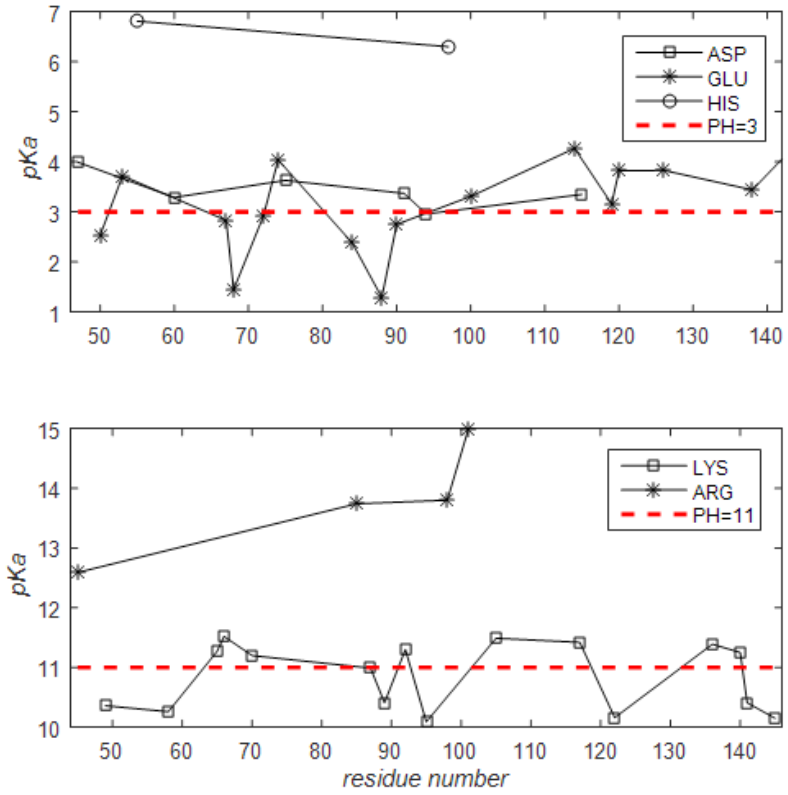

Fig. 4. pKa of various amino acids in acidic and alkali modes

Molecular dynamic simulation is accomplished for a chain of protein for two different sizes of the box. In simulations, the distance between the wall and mass center of protein is set to 1nm for one step, and 2nm for another step. Simulations are so that the number of the molecules of solvent are adjusted automatically by increasing the number of boxes. Indeed, increasing the value of solvent is to the extent that the density of solvent due to the effect of increasing the size of the box is constant.

Increasing the level of solvent leads to removing many terms regarding total energy. The higher size of box leads to increasing the level of solvent. Therefore, by creating boundary condition around the box, due to increasing distance, many intervening terms are created, lowered or removed among a chain of protein and adjacent proteins (Fig. 5).

Figs. 6 and 7 show displacing mass center and arm for the each chain of the protein against environmental condition change. According to the figures, it is evident that reducing temperature to 270 K leads to shrinking the arm and acidic environment causes the extension of the arm of a chain of Chaperone protein. Also, alkali environment causes oscillating behavior in the arm. On the other hand, an increase in the size of the box smoothens and stabilizes the behavior of the system. Therefore, temperature and pH are control inputs for a chain and facilitate condition for taking and releasing nano drug.

Nanorobots and bionano robots must be at least able to accomplish one of the characteristics such as stimulation, measurement, data processing, reaction or group behavior in nanoscale [21]. Results coming from molecular dynamic

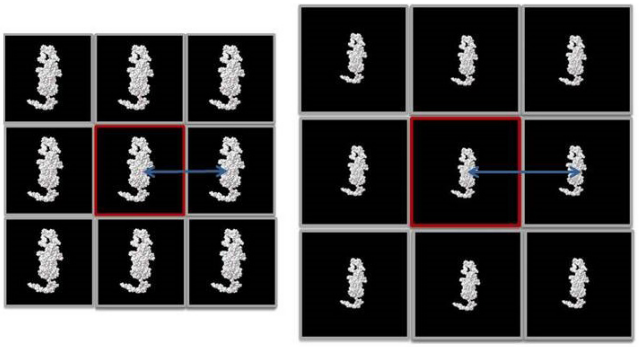

Fig. 5. Providing boundary condition around a chain of protein in molecular dynamic simulation in two boxes with different sizes.

simulation show that protein chains react well against external motives, and inputs cause a change in the behavior of the arm and mass center. Therefore, each chain can be regarded as a bio nano robot.

4- Modeling a Chain of Protein

Molecular dynamic simulation is time-consuming due to the high load of calculations, and providing a model of system behavior facilitates utilizing protein as a bionano robot in a shorter time. There are various methods to identify non-linear systems. These methods include Volter series models, block-oriented models, neuro network models and NARMAX. In this article, neural network models are used, with the following details.

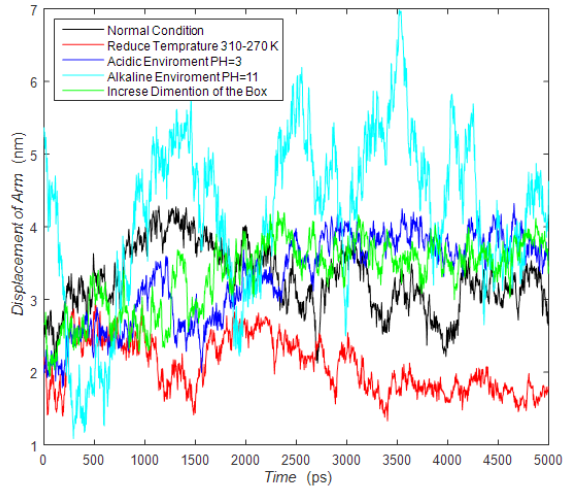

Fig. 6. Relocation of the arm for a chain of protein by changing ambient condition in molecular dynamic simulation

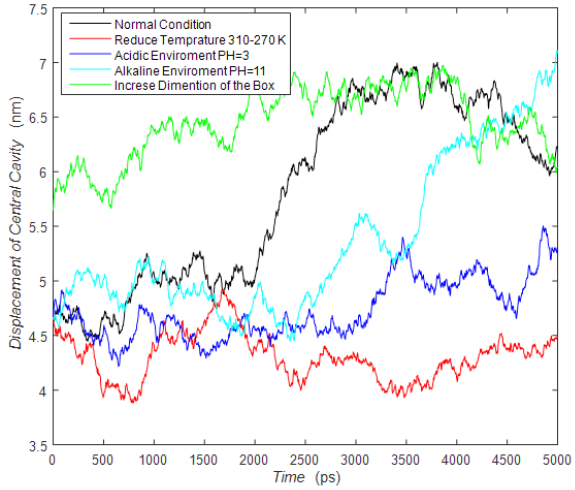

Fig. 7. Relocation of mass center for a chain of protein by changing ambient condition in molecular dynamic simulation

4- 1- Multi-layer perceptron

Multi-Layer Perceptron (MLP) with an intermediary layer and enough number of neurons can estimate a function with proper accuracy [22]. Input and output data are normalized to the neural network. Choosing initial weights are so as to be set in a linear area of stimulation function. The neural network has a hidden layer with a sigmoid stimulation function. This causes enough large gradient and learning process to be accomplished with more speed [23]. According to Fig. 8, it is observed that identification of errors and output correlation plot are like white noise. This shows that identification is done with a good accuracy.

4- 2- Radial basis function

Radial Basis Function (RBF) contains 3 layers, and its advantage to multi-layer perceptron is that it can identify non-linear systems with fewer numbers of layers and neurons [24]. The output is in linear form. Thus, it can be estimated by methods such as least square. According to Fig. 9, it is evident that identification of error has Gaussian with a mean near zero and acts like white noise.

4- 3- Adaptive neuro fuzzy inference system

One of the advantages of fuzzy set, in spite of its limitations, is the ability of combination with an artificial neural network which provides the condition for learning. Input membership functions of protein bio-nano robots are considered as Gaussian-bell type. Also, an optimization algorithm is considered as a combination of least square and back propagation, according to Fig. 10, it is observed that identification of error has decreased in comparison to previous modes and error correlation plot is like white noise and there is no dynamics in the error.

5- Results and Discussion

In section 4, three non-linear identifiers are used to identify a chain of sHSP. To compare the identification methods, four criteria are utilized. The criteria include mean absolute error, mean square error, mean absolute percentage error and root mean square error, defined as following equations:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \quad (1)$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \quad (3)$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (4)$$

where

Number of data	N
Real output, nm	y
Simulated output, nm	$\hat{y}$

According to Tables 1 and 2, it is shown that the margin of error is negligible in three methods. However, the least value of error in identification of arm and mass center displacement is for ANFIS. ANFIS provides the best model due to utilizing verbal variables in fuzzy approach and ability to learn with a neural network. Also, input data are so that we can describe a system with a simpler explanation using verbal variables, hence it is to say that if temperature decreases and the environment is null, then the arm of bio nano robot shrinks, or if the temperature is in normal mode and pH decreases, then arm extends.

The identification of data is obtained through molecular dynamic simulation under different conditions. Molecular dynamic simulation is time consuming because of high load of calculations, in a way that requires a 32-core computer for a week for this purpose. Nevertheless, a load of calculations regarding protein size increases. Therefore, by providing a valid model of system behavior, it is possible to accomplish analyses in a shorter time and provide a room to control a chain of proteins to take and release nano drugs.

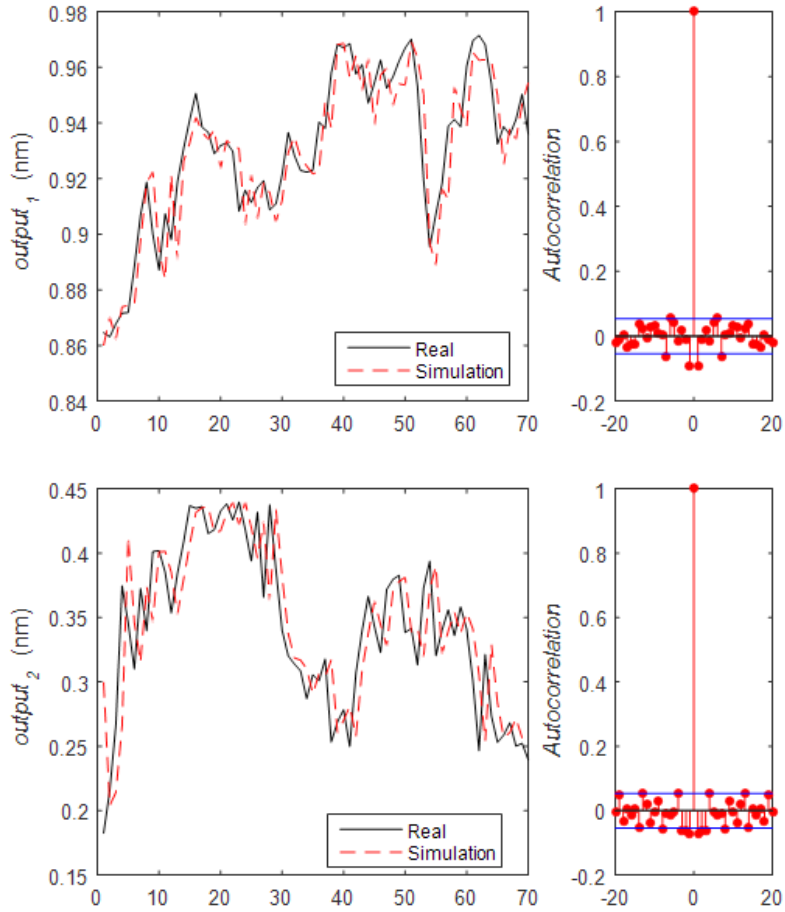

Fig. 8. Error correlation plot and identification behavior of a chain of sHSP based on MLP

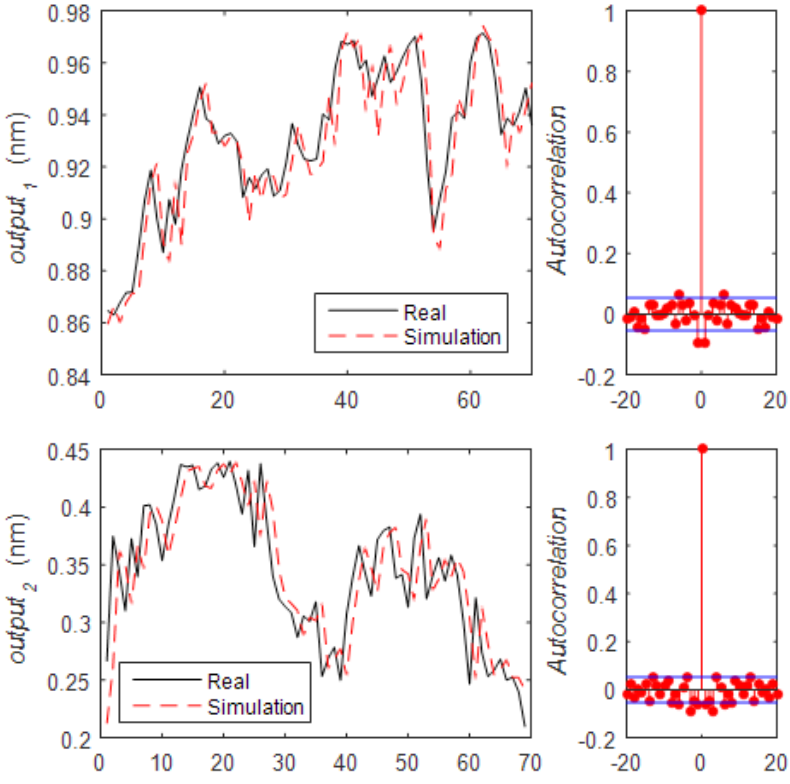

Fig. 9. Error correlation plot and identification behavior of a chain of sHSP based on radial basis function

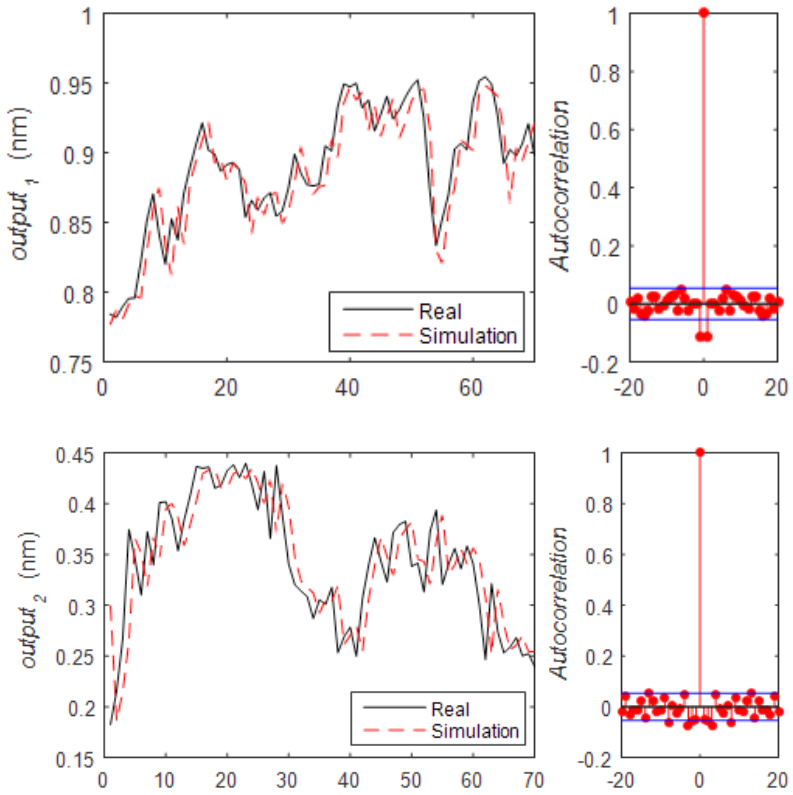

Fig. 10. Error correlation plot and identification behavior of a chain of sHSP based on ANFIS

Table 1. Comparison between non-linear identifier approaches based on various criteria for first output (mass center relocation)

Criteria	MLP	RBF	ANFIS
MAE	0.0096	0.0096	0.0087
MSE	0.0008	0.00079	0.00032
MAPE	1.7	1.69	1.54
RMSE	0.0277	0.0283	0.0179

Table 2. Comparison between non-linear identifier approaches based on various criteria for first output (arm relocation)

Criteria	MLP	RBF	ANFIS
MAE	0.0322	0.0312	0.0307
MSE	0.0021	0.0019	0.0017
MAPE	8.832	8.56	8.42
RMSE	0.045	0.043	0.0414

6- Conclusion

In this article, the behavior of a chain of sHSP Chaperone protein is examined based on molecular dynamic simulation. A chain of protein was introduced as bionano robot due to being agitated against external inputs and because of its appropriate biological structure. By applying inputs in a wide range of changes and obtaining data reach of system behavior, the behavior of arm and mass center for a chain of protein was modeled by the help of non-linear identifiers such as multi-layer perceptron, Radial basis function and adaptive neuro fuzzy inference system. According to results obtained, it was observed that ANFIS provides the least values of errors due to benefiting both advantages of fuzzy and neural networks as well as the type of data. Providing a dynamic model of a chain of sHSP facilitates the behavior analysis and control of the arm to utilize protein bio-nanorobots.

References

[1] R.J.K. T. Hogg, Mobile micro scopic sensors for high resolution in vivo diagnostics, Nano medicine: Nano technology, *Biology and Medicine*, 2(4) (2006) 239-247.

[2] A. Requicha, Instrumented cellular systems, Nanomedicine: Nanotechnology, *Biology and Medicine*, 2(4) (2006) 278-287.

[3] B.S. A. Calvacanti, R. Freitas,T. Hogg, Nanorobot architecture for medical target identification, *Nanotechnology*, 19(1) (2008) 1-15.

[4] A.F. M. Hamdi, DNA nanorobotics, *Microelectronics Journal*, 39(8) (2008) 1051-1059.

[5] C.B.K. E. Steager, J. Patel, S. Bith,C. Naik, L. Reber,M. J. Kim Control of microfabricated structures powered by flagellated bacteria using phototaxis, *Applied Physics Letters*, 90(26) (2006) 263901-263903.

[6] B.J.N. L.X. Dong, T. Fukuda, F. Arai, 2006. "Towards linear nano servomotors". , 3(3), pp. 228–235., , *IEEE Transactions on Automation Science and Engineering*, 3(3) (2006) 228-235.

[7] L.X.D. A. Subramanian, B. J. Nelson, Batch fabrication of carbon nanotube bearings, *Nanotechnology*, 18(7) (2007) 075703.

[8] K.K. G.S. Chirikjian, C. Mavroidis Analysis and design of protein based nanodevices: challenges and opportunities

in mechanical design, *Journal of Mechanical Design*, 127(4) (2005) 695–698.

[9] O.M.A. R. T. Richardson, G. Grossman, et al, Nuclear Autoantigenic Sperm Protein (NASP), a Linker Histone Chaperone That is required for Cell Proliferation, *Journal of Biological Chemistry*, 281(30) (2006) 21526–21534.

[10] S.M.v.d.V. R. J. Ellis, Molecular Chaperones, *Annual review of biochemistry*, 60(1) (1991) 321–347.

[11] A.S. A. Ghaffari, R. Hasanzadeh Ghasemi, Capturing and releasing a nano cargo by Prefoldin nano actuator, *Sensors and Actuators B: Chemical*, 171-172(1) (2012) 1199-1206.

[12] A.S. A. Ghaffari, R. Hasanzadeh Ghasemi, Design and simulation of a novel bio nano actuator by Prefoldin, in: *10th IEEE Conference on Nanotechnology*, 2010, pp. 885-888.

[13] A.S. A. Ghaffari, R. Hasanzadeh Ghasemi, Prefoldin: A nano actuator for carrying the various size nano drugs, *Journal of Computational and Theoretical Nanoscience*, 8(10) (2011) 2078-2086.

[14] C.S. R. Van Montfort, E. Vierling, Structure and function of the small heat shock protein /α-crystallin family of molecular chaperones, *Advances in Protein Chemistry* 59(1) (2001) 105–156.

[15] F. Narberhaus, Mechanism of chaperone function in small heat shock proteins, *Microbiology and Molecular Biology Reviews*, 66(1) (2002) 64–93.

[16] J.L.P.B. C. K. Kennaway, U. Gohlke, L. Wang, E. V. rlova, C. V. Robinson, H. R. Saibil, N. H. Keep, Dodecameric structure of the small heat shock protein Acr1 from Mycobacterium tuberculosis, *Journal of Biological Chemistry*, 280(39) (2005) 33419–33425.

[17] D. Shugar, Dynamics of Proteins and Nucleic Acids, *FEBS Letters*, 253(1) (1987) 248.

[18] D.v.d.S. B. Hess, E. Lindahl, et al, *GROMACS USER MANUAL Version 4.5.6, Department of Biophysical Chemistry*, University of Groningen & 2001–2010: The GROMACS development teams at the Royal Institute of Technology and Uppsala University, Sweden, 1991-2000.

[19] A.E.M. T. A. Wassenaar, The Effect of Box Shape on the Dynamic Properties of Proteins Simulated under Periodic Boundary Conditions, *Journal of computational chemistry*, 27(3) (2006) 316-325.

[20] K. Pauwels, Chaperoning Anfinsen: The Steric Foldases, *Molecular Microbiology*, 64(4) (2007) 917-922.

[21] G. Dorffner, Neural Networks for Time Series Processing, *Neural Network World*, 6(4) (1996) 447-468.

[22] G.B.O. G. Montavon, K. R. Muller, *Neural Networks: Tricks of the Trade*, Berlin, Springer., 2012.

[23] S.A.B. S. Chen, Neural network for nonlinear dynamic system modeling and identification, *International Journal of Control*, 56(2) (1992) 319–346.

Please cite this article using:

A. Gaffari, A. Khodayari, Sh. Movasagh, "Modeling Behavior of a Chain of Protein as a Bionanorobot by Changing Environmental Condition", *AUT J. Mech. Eng.*, 1(1) (2017) 109-116.

DOI: 10.22060/mej.2016.757